

&lt;!--StartFragment--&gt;RESULT 4

AB007453

LOCUS AB007453 3500 bp mRNA linear VRT 08-JUN-2002

DEFINITION Chrysophrys major mRNA for estrogen receptor, complete cds.

ACCESSION AB007453

VERSION AB007453.1 GI:2447037

KEYWORDS estrogen receptor.

SOURCE Pagrus major (red seabream)

ORGANISM Pagrus major

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;  
 Percoidei; Sparidae; Pagrus.

REFERENCE 1

AUTHORS Touhata,K., Kinoshita,M., Toyohara,H. and Sakaguchi,M.

TITLE Sequence and expression of a cDNA encoding the red seabream estrogen receptor

JOURNAL Fisheries Sci. 64, 131-135 (1998)

REFERENCE 2 (bases 1 to 3500)

AUTHORS Touhata,K. and Toyohara,H.

TITLE Direct Submission

JOURNAL Submitted (19-SEP-1997) Ken Touhata, Japan Society for the  
 Promotion Science, National Research Institute of Fisheries  
 Science; 2-12-4 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236-8648,  
 Japan (E-mail:ktouhata@fra.affrc.go.jp, Tel:81-45-788-7685,  
 Fax:81-45-788-5001)

FEATURES

source Location/Qualifiers

1. .3500

/organism="Pagrus major"

/mol\_type="mRNA"

/db\_xref="taxon:143350"

/tissue\_type="liver"

/note="Pagrus major"

CDS

207. .1952

/codon\_start=1

/product="estrogen receptor"

/protein\_id="BAA22517.1"

/db\_xref="GI:2447038"

/translation="MYPEDSRGSGGVATVDFLEGTYDYAAPTPAPTPLYSHSTPGYYS  
 APLDAHGPPSDGSLQSLGSGPNSPLVFPSSPRLSPFMHPPTHYLETSTPVYRSSV  
 PSSQQSVSREDQCGTSDDSYVSGESGAGALAAGFEIAKEMRFAVCSDYASGYHYGVW  
 SCEGCKAFFKRSIQGHNDYMC PATNQCTIDNRNRKSCQACRLRKCYEVMKGMGRKD  
 RGRVLRDRDKQRTGTS DRDKASKGLEHRTAPPQDRRKHISSAGGGGKSSMISMPDQ  
 VLLLLQGAEPMLCSRQKLNRPYTEVTMMTLLTSMADKELVHMIWAKKLPGLQLSL  
 HDQVQLLESSWLEVLMI GLIWRSIHCPGKLIFAQDLILDRSEGDCVEGMAEIFDMLLA  
 TASRFRMLKLKPEEFVCLKAIILLNSGAFSFGTGTMEPLHDGA AVQNMLDTITDALIH  
 HINQSGCSAQQQSRQAQLLLLLSHIRHMSNKGMEHLYSMKCKNKVPLYDLLLEMLDA  
 HRIHRADRP AETWSQADREPPFTSRNSSGGGGGGGGSSSAGSTSGPRVSHESPTSPG  
 VLQYGGSRSECTHIL"

polyA\_site

3500

/note="25 A nucleotides"

ORIGIN

Query Match 80.0%; Score 1396.8; DB 11; Length 3500;

Best Local Similarity 89.0%; Pred. No. 0;

Matches 1564; Conservative 0; Mismatches 167; Indels 27; Gaps 4;

Qy 1 ATGTACCCTGAAGAGAGCAGGGGGTCCGGAGGGGTAGCCACTGTGGACTTTCTGGAAGGG 60

||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 207 ATGTACCCGAAGACAGCCGGGGTCTGGAGGGGTAGCCACTGTGGACTTCTGGAAGGG 266

Qy	61	ACCTACGATTATGCCGCCCCCACCCTGCCCCGACTCCTCTTTACAGCCAGT-----CT	114
Db	267	ACGTACGACTATGCCGCCCCCACCCTGCCCCGACTCCTCTTTATAGCCACTCCACCCCT	326
Qy	115	GGCTACTACTCTGTACCTCTGGACGCCCCAAGGGCCACCCTCAGATGGCAGCCTTCAGTCC	174
Db	327	GGCTACTACTCTGTCTCCTCTGGACGCCCCACGGACCACCCTCAGATGGGAGCCTTCAGTCC	386
Qy	175	CTGGGCAGCGGGCCTACCAGTCCTCTTGTGTTTGTGCCGTCCAGCCCCAGACTCAGCCCC	234
Db	387	CTGGGCAGCGGGCCCAACAGTCCTCTAGTGTTTGTGCCCTCCAGCCCCCGTCTTAGCCCC	446
Qy	235	TTTATGCACCCGCCCAGCCACCCTATCTGGAAACCACCTCAACACCCGTCTACAGGTCT	294
Db	447	TTCATGCACCCGCCCACCACCCTATCTGGAAACCACCTCAACACCCGTCTACAGGTCC	506
Qy	295	AGTGTCCCATCCAGCCAGCAGCCAGTTCCAGAGAGGACCAGTGTGCCACCAGTGATGAG	354
Db	507	AGTGTCCCATCCAGTCAGCAGTCAGTTTCAAGAGAGGACCAGTGTGGCACCAGTGACGAC	566
Qy	355	TCCTATAGTGTGGGGGAGTCAGGGGCTGG-----AGCCAGGGGGTTTGAGATGGCCAAG	408
Db	567	TCATACAGCGTGGGAGAGTCAGGTGCCGGAGCGTTAGCGGCTGGGTTTGAGATTGCCAAA	626
Qy	409	GAGATGCGTTTTCTGTGCTGTGTGCAGTGACTATGCCTCTGGGTACCACTACGGGGTGTGG	468
Db	627	GAAATGCGTTTTCTGTGCCGTGTGCAGCGATTATGCCTCTGGGTACCACTACGGGGTGTGG	686
Qy	469	TCCTGTGAAGGCTGTAAGGCCTTCTTTAAGAGGAGCATCCAGGGTCACAATGACTATATG	528
Db	687	TCCTGTGAGGGCTGCAAGGCCTTCTTTAAGAGGAGCATACAGGGTCACAATGACTATATG	746
Qy	529	TGCCCAGCAACCAATCAGTGTACTATTGACAGGAATCGGAGAAAGAGCTGCCAGGCCTTGC	588
Db	747	TGCCCAGCAACCAATCAGTGTACTATTGACAGGAATCGGAGGAAGAGCTGCCAGGCCTTGC	806
Qy	589	CGTCTTAGGAAGTGTTATGAAGTGGGCATGATGAAAGGAGGTGTTGCAAGGACCGTGGC	648
Db	807	CGTCTTAGGAAGTGTTACGAAGTGGGCATGATGAAAGGAGGTATGCGCAAGGACCGTGGG	866
Qy	649	CGTGTTTTGCGCCGTGATAAACGACGTGCTGGAACCAATGACCGAGAGAAGGCCTCTAAG	708
Db	867	CGCGTTCTGCGGCGTGACAAACAACGGACTGGCACCAGTGACAGAGACAAGGCCTCTAAG	926
Qy	709	GACCTGGAGTACAAAACAGTGCCCCCTCAGGACAGGAGGAAACACAGCAGCAGCAGT	768
Db	927	GGTCTGGAGCACAGAACAGCGCCCCCTCAGGACAGGAGGAAACA---CATCAGCAGCAGT	983
Qy	769	GCCGGTGGTGGAGGAGGAAAATCATCAGTGACCGGGATGTCTCCTGACCAGGTGCTCCTC	828
Db	984	GCTGGTGGTGGTGGAGGAAAGTCATCAATGATCAGCATGCCTCCTGACCAGGTGCTCCTC	1043
Qy	829	CTGCTCCAGGGTGCCGAGCCCCCAATGCTGTGCTCCCGTCAGAAGCTGAGCCGACCGTAC	888
Db	1044	CTGCTCCAGGGTGCCAGAGCCCCCAATGCTGTGCTCCCGTCAGAAGCTGAACCGACCCCTAC	1103
Qy	889	ACCGAGGTCACCATAATGACACTACTCACCAGCATGGCCGATAAGGAGCTGGTCCACATG	948
Db	1104	ACCGAGGTCACCATGATGACCCTGCTCACCAGCATGGCTGATAAGGAGCTGGTCCACATG	1163
Qy	949	ATCACCTGGGCCAAGAAGCTTCCAGGTTTCTGTCAGCTGTCTCTCCATGACCAGGTGCAG	1008

```

Db      1164  ATCGCATGGGCAAAGAAGCTTCCAGGTTTCCTACAGCTGTCCCTCCACGACCAGGTGCAG 1223
Qy      1009  CTGCTGGAGAGCTCGTGGCTGGAGGTGCTGATGATTGGGCTCATATGGAGGTCCATCCAC 1068
Db      1224  CTGCTGGAGAGCTCGTGGCTGGAGGTGCTGATGATCGGGCTCATCTGGAGGTCCATCCAC 1283
Qy      1069  TGCCCCGGCAAACCTCATCTTCGCACAGGACCTCATACTGGACAGGAATGAAGGTGACTGT 1128
Db      1284  TGCCCCGGCAAACCTCATCTTCGCACAGGACCTCATACTGGACAGGAGCGAAGGCGACTGT 1343
Qy      1129  GTGGAAGGCTTTGTTGAGATCTTCGACATGCTGCTGGCCACTGCCTCCCGCTTCCGCATG 1188
Db      1344  GTCGAAGGCATGGCTGAGATCTTCGACATGCTGCTTGCCACCGCTCCCGTTTCCGCATG 1403
Qy      1189  CTCAAACCTCAAACCTGAGGAGTTTGTCTGCCTCAAAGCTATCATCCTGCTCAACTCTGGT 1248
Db      1404  CTCAAACCTCAAACCTGAGGAGTTCGTCTGCCTCAAAGCTATCATCTTGCTCAACTCTGGT 1463
Qy      1249  GCCTTCTCTTTCTGCACCGGCACAATGGAGCCCTCCACAACAGCATGGCAGTGCAGAAC 1308
Db      1464  GCCTTCTCTTTCTGCACCGGCACAATGGAGCCCTCCACGACGGTGCAGCAGTGCAGAAC 1523
Qy      1309  ATGCTGGACACCATCACAGACGCTCTCATAATCATATCAGCCAATCAGGATGCTCGGCT 1368
Db      1524  ATGCTCGACACCATCACCGACGCTCTCATAATCATATCAACCAATCCGGATGCTCGGCT 1583
Qy      1369  CAGCAGCAGTCGAGGCGGCAGGCCAGCTGCTGCTCCTGCTCTCCACATCAGGCACATG 1428
Db      1584  CAGCAGCAGTCGAGGCGGCAGGCCAGCTGCTCCTGCTCTCCACATCAGACACATG 1643
Qy      1429  AGCAACAAAGGCATGGAGCATCTCTACAGCATGAAGTGCAAGAACAAGTGCCTCTTTAC 1488
Db      1644  AGCAACAAAGGCATGGAGCATCTCTACAGCATGAAGTGCAAGAACAAGTGCCTCTGTAC 1703
Qy      1489  GACCTTCTGCTGGAGATGTTGGACGCTCACCGTATACACCGCCCAGACAGACCAGCTCAG 1548
Db      1704  GACCTGCTGCTGGAGATGCTGGACGCTCACCGCATCCACCGCGCCGACAGACCAGCTGAG 1763
Qy      1549  TTCTGGTCCCAGGCTGACGGAGAGCCTCCCTTCATTAACAACAACAACAGCAGCAACAGT 1608
Db      1764  ACCTGGTCCCAGGCTGACAGAGAGCCTCCCTTCACCTCCAGAAACAGCAGCGGTGGTGGT 1823
Qy      1609  GGCAGCAATGGCGGCGTCTCCTCTTCAGTCGGTTCCAGTTCAGGACCCCGAGTCAACCAC 1668
Db      1824  GGTGGTGGTGGAGGAGGCTCCTCTTCAGCTGGCTCCACTTCAGGACCACGAGTCAGCCAT 1883
Qy      1669  GAGAGCCCGAGCAGAGGACCCACAGGTCCAGGAGTCCTGCAGTACGGAGGGTCCCGCTCT 1728
Db      1884  GAGAGC-----CCCACAAGTCCTGGTGTCTGTCAGTACGGAGGGTCCCGCTCC 1931
Qy      1729  GACTGCACCCACATCCTA 1746
Db      1932  GAGTGCACTCACATCCTA 1949
```

<!--EndFragment-->

```

<!--StartFragment-->RESULT 3
ESR1_PAGMA
ID   ESR1_PAGMA      STANDARD;          PRT;   581 AA.
AC   O42132;
DT   15-JUL-1998, integrated into UniProtKB/Swiss-Prot.
DT   01-JAN-1998, sequence version 1.
DT   30-MAY-2006, entry version 47.
DE   Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).
GN   Name=ESR1; Synonyms=ESR, NR3A1;
OS   Pagrus major (Red sea bream) (Chrysophrys major).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC   Sparidae; Pagrus.
OX   NCBI_TaxID=143350;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA].
RC   TISSUE=Liver;
RA   Touhata K., Kinoshita M., Toyohara H., Sakaguchi M.;
RT   "Sequence and expression of a cDNA encoding the red seabream estrogen
RT   receptor.";
RL   Fish. Sci. 64:131-135(1998).
CC   -!- FUNCTION: The steroid hormones and their receptors are involved in
CC   the regulation of eukaryotic gene expression and affect cellular
CC   proliferation and differentiation in target tissues.
CC   -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with ER-
CC   beta (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nucleus.
CC   -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC   a DNA-binding domain and a C-terminal steroid-binding domain.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC   subfamily.
CC   -!- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
CC   -----
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----
DR   EMBL; AB007453; BAA22517.1; -; mRNA.
DR   HSSP; P03372; 1ERR.
DR   SMR; O42132; 140-212.
DR   InterPro; IPR001628; Hrmn_rcpt_DNA_bd.
DR   InterPro; IPR000536; Hrmn_rcpt_lig_bd.
DR   InterPro; IPR008946; Nucl_rcpt_lig_bd.
DR   InterPro; IPR001292; Oestr_rcpt_AF1.
DR   InterPro; IPR012239; Oestrogen_rcpt.
DR   InterPro; IPR001723; Str_hrmn_rcpt.
DR   InterPro; IPR000324; VitD_rcpt.
DR   Pfam; PF00104; Hormone_recep; 1.
DR   Pfam; PF02159; Oest_recep; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PIRSF; PIRSF002527; ER-ab; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   PRINTS; PR00350; VITAMINDR.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; Znf_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
DR   PROSITE; PS51030; NUCLEAR_REC_DBD_2; 1.
KW   DNA-binding; Lipid-binding; Metal-binding; Nuclear protein; Receptor;
KW   Steroid-binding; Transcription; Transcription regulation; Zinc;

```

KW	Zinc-finger.			
FT	CHAIN	1	581	Estrogen receptor.
FT				/FTid=PRO_0000053635.
FT	DNA_BIND	145	210	Nuclear receptor.
FT	ZN_FING	145	165	NR C4-type.
FT	ZN_FING	181	205	NR C4-type.
FT	REGION	1	144	Modulating (By similarity).
FT	REGION	211	272	Hinge.
FT	REGION	273	581	Steroid-binding (By similarity).
FT	COMPBIAS	261	265	Poly-Gly.
FT	COMPBIAS	278	281	Poly-Leu.
FT	COMPBIAS	469	473	Poly-Leu.
FT	COMPBIAS	537	545	Poly-Gly.
SQ	SEQUENCE	581 AA; 63461 MW; D499BB6E7B4EB964 CRC64;		

Query Match 89.1%; Score 2752.5; DB 1; Length 581;  
 Best Local Similarity 89.1%; Pred. No. 1.7e-189;  
 Matches 522; Conservative 22; Mismatches 33; Indels 9; Gaps 4;

Qy	1	MYPEESRSGGVATVDFLEGTYDYAAPTAPTPLYSQS--GYYSVPLDAQPPSDGSLQS	58
Db	1	MYPEDSRSGGVATVDFLEGTYDYAAPTAPTPLYSHSTPGYYSAPLDAHGPPSDGSLQS	60
Qy	59	LGSGPTSPLVFPSSPRLSPFMHPPSHHYLETTSTPVYRSSVPSSQQVPVPREDQCATSDE	118
Db	61	LGSGPNSPLVFPSSPRLSPFMHPPTHHYLETTSTPVYRSSVPSSQQSVSREDQCGTSD	120
Qy	119	SYSVGESGAG--ARGFEMAKEMRFCVCSGYHYGVWSCEGCKAFFKRSIQGHNDYM	176
Db	121	SYSVGESGAGALAAGFEIAKEMRFCVCSGYHYGVWSCEGCKAFFKRSIQGHNDYM	180
Qy	177	CPATNQCTIDNRNRKSCQACRLRKCYEVGMMKGGVRKDRGRVLRRDKRRAGTNDREKASK	236
Db	181	CPATNQCTIDNRNRKSCQACRLRKCYEVGMMKGGMRKDRGRVLRRDKQRTGTSDRDKASK	240
Qy	237	DLEYKTVPPQDRRKHSSSSSAGGGGGKSSVTGMSPDQVLLLLQGAEPMLCSRQKLSRPY	296
Db	241	GLEHRTAPPQDRRKH-SSSSAGGGGGKSSMISMPDQVLLLLQGAEPMLCSRQKLNRPY	299
Qy	297	TEVTIMTLLTSMADKELVHMITWAKKLPGLQLSLHDQVQLLESSWLEVLMIGLIWRSIH	356
Db	300	TEVTMTLLTSMADKELVHMIWAKKLPGLQLSLHDQVQLLESSWLEVLMIGLIWRSIH	359
Qy	357	CPGKLIFAQDLILDRNEGDCVEGFVEIFDMLLATASRFRMLKLKPEEFVCLKAIILLNSG	416
Db	360	CPGKLIFAQDLILDRSEGDCVEGMAEIFDMLLATASRFRMLKLKPEEFVCLKAIILLNSG	419
Qy	417	AFSFACTGTMEPLHNSMAVQNMLDTITDALIHHSQSGCSAQQQSRRQAQLLLLLSHIRHM	476
Db	420	AFSFACTGTMEPLHDGAAVQNMLDTITDALIHHSQSGCSAQQQSRRQAQLLLLLSHIRHM	479
Qy	477	SNKGMEHLYSMKCKNKVPLYDLLLLLEMLDAHRIHRPDRPAQFWSQADGEPPFINNNSSNS	536
Db	480	SNKGMEHLYSMKCKNKVPLYDLLLLLEMLDAHRIHRADRPATWSQADREPPFTSRNSSGGG	539
Qy	537	GSNGGVSSSVGSSSGPRVNHESPSRGPTGPGVLQYGGSRSDCTHIL	582
Db	540	GGGGGGSSSAGSTSGPRVSHES----PTSPGVLQYGGSRSECTHIL	581

<!--EndFragment-->